

Code Generation: Results Achieved with Automated Chain-of-Thought prompting Considering Only Methods with No External Dependencies

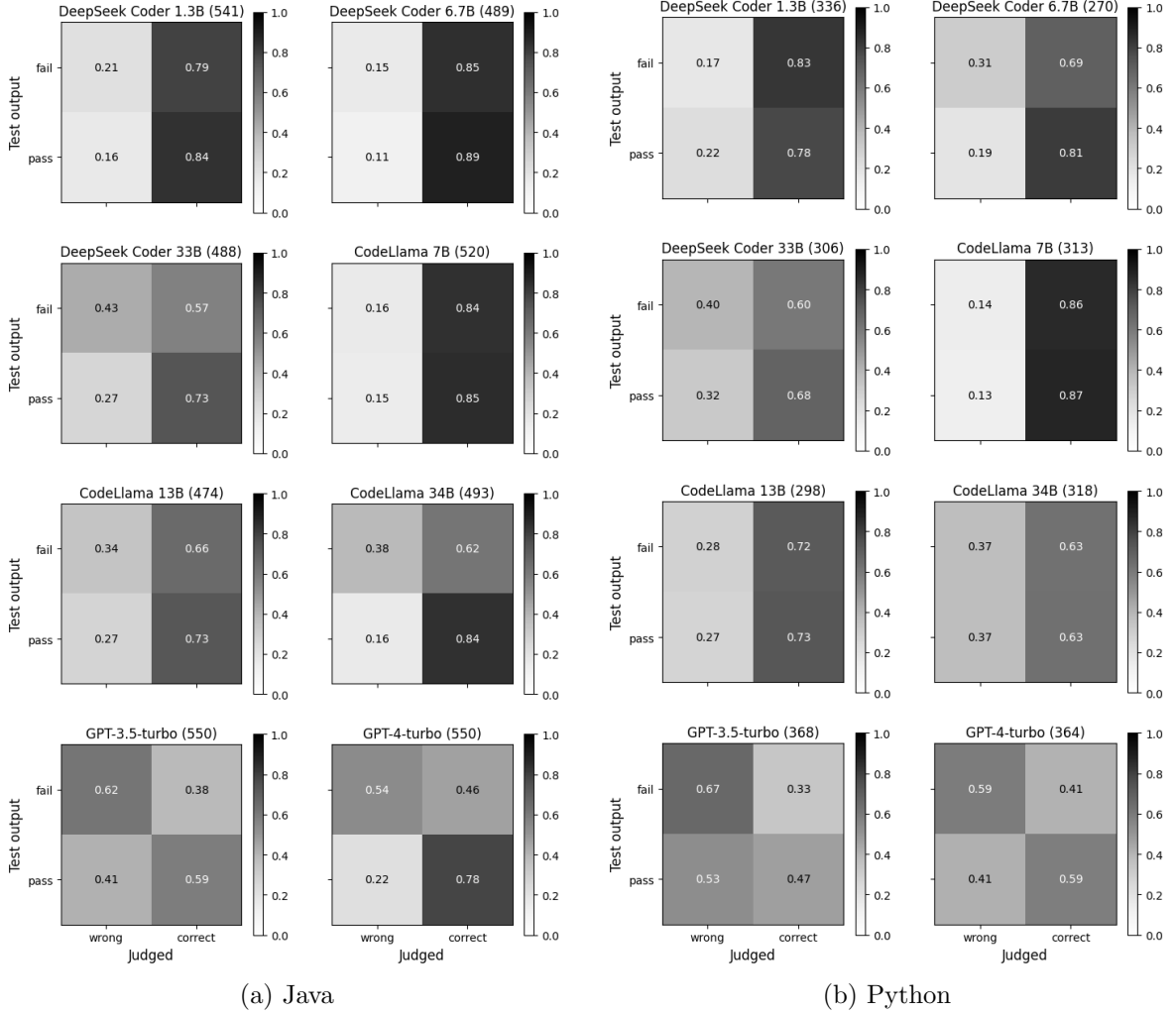


Figure 1: Code Generation: Confusion matrices for LLM’s judgment.

	DSC	DSC	DSC	CL	CL	CL	GPT	GPT
	1.3B	6.7B	33B	7B	13B	34B	3.5	4
Java	0.05	0.05	0.17	0.01	0.08	0.24	0.20	0.33
Python	-0.04	0.11	0.07	0.01	0.01	0.00	0.14	0.17

Table 1: Code Generation (only methods with no external dependencies): Kappa score.